JOURNAL OF FUNCTIONAL ANALYSIS **160**, 367–381 (1998)
ARTICLE NO. FU983290

Resolvent Estimates for Fleming–Viot Operators with Brownian Drift

Peter March*

The Ohio State University, Department of Mathematics, 231 West 18th Avenue, Columbus, Ohio 43210

Received March 22, 1996; revised May 2, 1998; accepted May 2, 1998

This article is a supplement to the paper of D. A. Dawson and P. March (*J. Funct. Anal.* **132** (1995), 417–472). We define a two-parameter scale of Banach spaces of functions defined on $\mathscr{M}_1(\mathbb{R}^d)$, the space of probability measures on d-dimensional euclidean space, using weighted sums of the classical Sobolev norms.

These estimates gauge the degree of smoothing performed by the resolvent and separate the contribution due to the diffusion coefficient and that due to the drift coefficient.

Key Words: Fleming–Viot process; resolvent; Sobolev space.

INTRODUCTION

Fleming–Viot processes form a class of probability measure-valued processes arising from a construction whose initial data are a Markov process on a state space X with generator $(A, \mathscr{A})$, $\mathscr{A} \subset C(X)$, and a nonnegative, symmetric function $\gamma: X \times X \to \mathbb{R}$. The result of the construction is a unique process ξ with values in $\mathscr{M}_1(X)$, the space of probability measures on X, having the characteristic properties that for any $\phi \in \mathscr{A}$,

$$M_t(\phi) = \xi_t(\phi) - \int_0^t \xi_s(A\phi)\, ds$$

is a martingale with quadratic variation

$$\langle M(\phi) \rangle_t = \int_0^t \xi_s^2(\gamma(\phi, \phi))\, ds$$

where

$$\gamma(\phi, \phi)(x, y) = \gamma(x, y)[\phi(x) - \phi(y)]^2$$

* Research supported by Grant DMS-9307706 from the National Foundation.

0022-1236/98 $25.00

and ξ^2 denotes two-fold product measure. Roughly speaking, ξ is constructed as the limiting empirical measure of a system of particles undergoing independent A-motions on which is superimposed a two-body interaction determined by γ. In the context of population genetics (cf. [EK1, EK2]) ξ is a diffusion approximation to the empirical distribution of a large population over a space of possible genetic types. The type of individual members of the population changes by *mutation*, which is characterized by A, and by *random mating*, which is characterized by γ. Thus it is apparent that mutation causes the empirical distribution to drift and random mating causes it to fluctuate. The result is a diffusion in the space of probabilities $\mathcal{M}_1(X)$ whose generator $\mathcal{L}$ can be written concretely as a differential operator in terms of the directional derivative

$$\nabla_x F(\mu) = \lim_{\varepsilon \downarrow 0} \frac{F(\varepsilon\delta + (1-\varepsilon)\,\mu) - F(\mu)}{\varepsilon}.$$

Indeed, for suitable functions F we have (cf. [DM])

$$\mathcal{L}F(\mu) = \tfrac{1}{2} \int_{X^2} \gamma(x, y)(\nabla_x - \nabla_y)^2\, F(\mu)\, \mu(dx)\, \mu(dy) + \int_X A\nabla_x F(\mu)\, \mu(dx).$$

For further background consult the surveys of Ethier and Kurtz [EK2] and Dawson [D].

The details of the construction were first carried out by Fleming and Viot [FV] in the case where X is compact, γ is constant, and A is *triangularisable*, following work on the case where X is a finite set and therefore ξ is a diffusion on a finite-dimensional simplex. See [E, F, Sa, Sh], for example. There is a sophisticated version due to Donnelly and Kurtz [DK] in the case where X is a locally compact, separable, metric space, $(A, \mathcal{A})$ generates a Feller process, and γ is constant. Versions of the construction for general γ can be found in Dawson [D, Sect. 5.7.8], and in Dawson and March [DM], Section 4.

In any event, ξ is a natural process determined by two parameters, the drift coefficient $(A, \mathcal{A})$ and the diffusion coefficient γ. In Dawson and March [DM] these processes were used as local models to construct measure-valued processes whose drift and diffusion coefficients are permitted to depend on the current state of the process; that is, conditionally on $\xi_t = \mu$, the process looks infinitesimally like a Fleming–Viot process with drift $(A(\mu), \mathcal{A})$ and diffusion coefficient $\gamma(\mu, x, y)$. The proof of uniqueness of the process with given coefficients of a restricted type depended crucially on estimates of the resolvent $\mathcal{R}_\lambda := (\lambda - \mathcal{L})^{-1}$ of a Fleming–Viot process acting on a two-parameter scale of Banach spaces of continuous functions $\mathcal{C}_{a,b}(\mathcal{M}_1(X)) \subset C(\mathcal{M}_1(X))$. It should be remarked that the type of coefficient permitted in the uniqueness proof in [DM] was determined by the nature of these resolvent estimates.

The purpose of this paper is to derive new resolvent estimates in the special case where $X=\mathbb{R}^d$, $\gamma(x, y)\equiv\gamma$ is constant, and $A=\kappa\Delta$, $\kappa>0$. These estimates are most conveniently expressed by defining a new two-parameter scale of Banach spaces and showing that the resolvent, in this special case, acts boundedly between certain members of the scale. They are similar in nature to estimates in [DM] and so the results presented here can be thought of as a supplement to that paper. The chief distinction between these estimates and the previous ones is that they are phrased in terms of the classical Sobolev spaces $H^s((\mathbb{R}^d)^n)$, $n\geqslant 1$, which allows us to gauge the degree of smoothing performed by the drift coefficient $A=\kappa\Delta$ as well as smoothing performed by the diffusion coefficient. Clearly, this is an infinite-dimensional feature of our problem, as one does not expect the drift coefficient of a finite-dimensional diffusion to be responsible for any smoothing effect of the resolvent.

To be more precise, we must introduce some notation. Let

$$\mathscr{X}=\coprod_{n=0}^{\infty}(\mathbb{R}^d)^n$$

be the topological direct sum of the cartesian powers of $\mathbb{R}^d$, where $(\mathbb{R}^d)^0$ is by convention a singleton set. Let

$$\mathscr{S}(\mathscr{X})=\prod_{n=0}^{\infty}\mathscr{S}((\mathbb{R}^d)^n),$$

where $\mathscr{S}((\mathbb{R}^d)^n)$ denotes the Schwartz class functions and by convention, $\mathscr{S}((\mathbb{R}^d)^0)=\mathbb{R}$. Let

$$\mathscr{S}_c(\mathscr{X})=\{f=(f_0, f_1, ...)\in\mathscr{S}(\mathscr{X})\text{: } f_n\equiv 0 \text{ for all but finitely many } n\}$$

and define a map Λ: $\mathscr{S}_c(\mathscr{X})\to C(\mathscr{M}_1(\mathbb{R}^d))$ by the formula

$$\Lambda(f)(\mu)=\sum_{n=0}^{\infty}\langle f_n, \mu^n\rangle,$$

where

$$\langle f_n, \mu^n\rangle=\int_{(\mathbb{R}^d)^n} f_n(x_1, ..., x_n)\,\mu(dx_1)\cdots\mu(dx_n),$$

and $\langle f_0, \mu^0\rangle=f_0$. The range of Λ, denoted $P^{\mathscr{S}}(\mathscr{M}_1(X))$, is the set of polynomial functions on $\mathscr{M}_1(X)$ with Schwartz class coefficients and it serves as a class of test functions. One can define various norms on $P^{\mathscr{S}}(\mathscr{M}_1(X))$ and then complete the polynomials to get Banach spaces of continuous functions. The recipe we use is as follows.

For each n let $\|\cdot\|_{(n)}$ stand for a norm on $\mathscr{S}((\mathbb{R}^d)^n)$ finer than the supremum norm; that is, there exists a constant C_n such that for all $n \geqslant 1$,

$$\|f_n\|_{C_0((\mathbb{R}^d)^n)} \leqslant C_n \|f_n\|_{(n)}.$$

Let $\rho = (\rho_0, \rho_1, \rho_2, ...)$ be a sequence of positive real numbers such that $\sup_n (C_n/\rho_n) < \infty$. It is not hard to prove that

$$\|F\|_\rho := \inf \left\{ \sum_{n=0}^{\infty} \rho_n \|f_n\|_{(n)} : f \in \mathscr{S}_c(\mathscr{X}) \text{ and } F = \Lambda(f) \right\}$$

defines a norm on $P^{\mathscr{S}}(\mathscr{M}_1(\mathbb{R}^d))$ finer than the supremum norm. (Note that typically, $F \in P^{\mathscr{S}}(\mathscr{M}_1(X))$ does not have a unique representation as $\Lambda(f)$ for some $f \in \mathscr{S}_c(\mathscr{X})$; that is, Λ is not one-to-one.) It is also not hard to show that the completion of polynomials with respect to this norm consists of all functions of the form

$$F(\mu) = \sum_{n=0}^{\infty} \langle f_n, \mu^n \rangle,$$

where each f_n is in the completion of $\mathscr{S}((\mathbb{R}^d)^n)$ relative to the norm $\|\cdot\|_{(n)}$ and $\sum_{n=0}^{\infty} \rho_n \|f_n\|_{(n)} < \infty$. The arguments are straightforward, if tedious, and essentially the same as in the proof of Proposition 2.D.2 in [DM].

The choice made in [DM] is $\rho_n = (1+n)^a b^n$ for $a \geqslant 0$, $b \geqslant 1$, and each $\|\cdot\|_{(n)}$ is the supremum norm on $C(X^n)$. If we use the notation $\|\cdot\|_{a,b}$ for the corresponding norm on polynomials, then the main estimate of that paper is the following: if $b > 1$ then there is a universal constant C such that

$$\|\mathscr{R}_\lambda F\|_{a+2,b} \leqslant C(\gamma^{-1} + \lambda^{-1})\, b(b-1)^{-1} \|f\|_{a,b},$$

where $\mathscr{R}_\lambda$ is the resolvent of the Fleming–Viot process. Note that for fixed $b > 1$ this estimate records an improvement by two degrees in a in the sense that the sup norm of the nth component of $\mathscr{R}_\lambda F$ is on average, with respect to the weights b^n, $n \geqslant 0$, no bigger than n^{-2} times the sup norm of the nth component of F. But this norm is incapable of recording any gain in smoothness of the nth component of $\mathscr{R}_\lambda F$ as a function of its n variables valued in $\mathbb{R}^d$.

One way to try to measure such a smoothing property of the resolvent is to make the following choices. Let $\|\cdot\|_{(n)}$ be the Sobolev norm of $H^s((\mathbb{R}^d)^n)$ for a suitable choice of s. By examining the constant in the

Sobolev embedding theorem, one can show that if $s \geqslant 1 + nd/2$ then the H^s norm is uniformly finer than the sup norm on $\mathscr{S}((\mathbb{R}^d)^n)$ and so is a suitable choice for the construction of a norm on $P^{\mathscr{S}}(\mathscr{M}_1(\mathbb{R}^d))$. Now we can state the main result of this paper.

THEOREM. *Let $\nu_n = 1 + nd/2$, $N_n = \sum_{j=1}^n \nu_j$, and*

$$\omega_n = \sqrt{\frac{2^{N_{n-1}}}{(16\pi)^{\nu_n - 1}\,\Gamma(\nu_n)}}.$$

For $-\infty < \alpha < \infty$ and $0 \leqslant \beta \leqslant 2$, the formula

$$\|F\|_{\alpha,\beta} = \inf\left\{ \sum_{n=0}^{\infty} (1+n)^{\alpha}\, \omega_n\, \|f_n\|_{H^{\nu_n+\beta}((\mathbb{R}^n)^n)} : f \in \mathscr{S}_c(\mathscr{X}) \text{ and } F = \Lambda(f) \right\}$$

defines a norm on polynomials $P^{\mathscr{S}}(\mathscr{M}_1(X))$. Let $\mathscr{H}^{\alpha,\beta}(\mathscr{M}_1(X))$ be the completion of $P^{\mathscr{S}}(\mathscr{M}_1(\mathbb{R}^d))$ relative to this norm.

Let $\mathscr{R}_\lambda$ denote the resolvent of the Fleming–Viot process with drift coefficient $A = \kappa\Delta$ and diffusion coefficient $\gamma(x, y) \equiv \gamma$, where $\kappa, \gamma > 0$.

(i) *If $\alpha + \beta < 1$, then $\mathscr{R}_\lambda$: $\mathscr{H}^{0,0} \to \mathscr{H}^{\alpha,\beta}$ and*

$$\|\mathscr{R}_\lambda F\|_{\alpha,\beta} \leqslant C\, \frac{(\kappa^{-1}+\lambda^{-1})^{\beta/2}\,(\gamma^{-1}+\lambda^{-1})^{1-\beta/2}}{1-(\alpha+\beta)}\, \|F\|_{0,0}.$$

(ii) *If $\alpha + \beta > 1$, then $\mathscr{R}_\lambda$: $\mathscr{H}^{\alpha+\beta-1,0} \to \mathscr{H}^{\alpha,\beta}$ and*

$$\|\mathscr{R}_\lambda F\|_{\alpha,\beta} \leqslant C\, \frac{(\kappa^{-1}+\lambda^{-1})^{\beta/2}\,(\gamma^{-1}+\lambda^{-1})^{1-\beta/2}}{\alpha+\beta-1}\, \|F\|_{\alpha+\beta-1,0}.$$

The constant C is independent of α, β, and λ.

Notice that in general there is a trade-off between the degree of smoothing achieved in the spatial variables $x \in (\mathbb{R}^d)^n$ for fixed n, indicated by parameter β, and the average decrease in size of the Sobolev norm of the nth component, indicated by parameter α. There are two special cases, however, in which α and β decouple. First, in item (i), if $\alpha = 0$, then

$$\|\mathscr{R}_\lambda F\|_{0,\beta} \leqslant C\, \frac{(\kappa^{-1}+\lambda^{-1})^{\beta/2}\,(\gamma^{-1}+\lambda^{-1})^{1-\beta/2}}{1-\beta}\, \|F\|_{0,0} \qquad 0 \leqslant \beta < 1.$$

Second, in item (ii), if $\beta = 1$, then

$$\|\mathscr{R}_\lambda F\|_{\alpha,1} \leqslant C\, \frac{(\kappa^{-1}+\lambda^{-1})^{1/2}\,(\gamma^{-1}+\lambda^{-1})^{1/2}}{\alpha}\, \|F\|_{\alpha,0} \qquad \alpha > 0.$$

These two inequalities show a sense in which, uniformly in n, there is a gain of up to one weak derivative in the spatial variables.

In finite dimensions one knows that the resolvent of the Laplacian yields a gain of two weak derivatives in the spatial variables and this amount of regularity is optimal. In the present case, however, to achieve the optimal gain of derivatives there must be a trade-off. Specifically, in item (ii), if $\beta = 2$, then

$$\|\mathscr{R}_\lambda F\|_{\alpha, 2} \leqslant C \frac{(\kappa^{-1} + \lambda^{-1})}{\alpha + 1} \|F\|_{\alpha+1, 0} \qquad \alpha > -1.$$

It is not clear that the spaces $\mathscr{H}^{\alpha, \beta}$ are the natural ones to use or that these estimates are in any sense sharp. However, they are first steps in the regularity theory of an interesting and natural operator, namely the generator $\mathscr{L}$ of the Fleming–Viot process with brownian drift. From the functional analytic point of view $\mathscr{L}$ is a somewhat singular object, an infinite-dimensional, degenerate elliptic differential operator, so one cannot expect sharp results to fall out easily.

Preparations. Let us recall a formula proved in [DM] for the resolvent acting on $\mathscr{P}^{\mathscr{S}}(\mathscr{M}_1(\mathbb{R}^d))$. By analysing this formula into familiar pieces and using well-known estimates for each piece we can establish the estimates announced in the theorem.

Let us first set down some useful notation. For each $n \geqslant 2$ and $1 \leqslant i \neq j \leqslant n$ define the linear transformation $s_{ij}^{n-1}: (\mathbb{R}^d)^{n-1} \to (\mathbb{R}^d)^n$ by

$$s_{ij}^{n-1}(x_1, ..., x_{n-1}) = (x_1, ..., x_{i \vee j - 1}, x_{i \wedge j}, x_{i \vee j}, ..., x_{n-1}).$$

Thus s_{ij}^{n-1} maps $(\mathbb{R}^d)^{n-1}$ onto the hyperplane

$$H_{ij}^{n-1} = \{ y = (y_1, ..., y_n) \in (\mathbb{R}^d)^n \colon y_i = y_j \}$$

of codimension d. It induces a map $S_{ij}^{n-1}: \mathscr{S}((\mathbb{R}^d)^n) \to \mathscr{S}(\mathbb{R}^d)^{n-1})$ by the rule

$$S_{ij}^{n-1} f_n = f_n \circ s_{ij}^{n-1}$$

and a map $S^{n-1}: \mathscr{S}((\mathbb{R}^d)^n) \to \mathscr{S}((\mathbb{R}^d)^{n-1})$ via summation:

$$S^{n-1} f_n = \sum_{i=1}^{n} \sum_{j \neq i}^{n} S_{ij}^{n-1} f_n.$$

If $(A, \mathscr{A})$ is the generator of a Markov process on $\mathbb{R}^d$, we use $(A_n, \mathscr{A}_n)$ to denote the generator of a vector of n independent processes each with generator $(A, \mathscr{A})$. For the time being we take $\gamma = 1$ and $A = \kappa \Delta$ since we can reduce to this case by a scaling argument.

The proof of the following proposition follows easily from the proof of [DM, Theorem 3.B.2].

PROPOSITION A. *Let $\mathscr{R}_\lambda = (\lambda - \mathscr{L})^{-1}$ be the resolvent of the Fleming–Viot process with $\gamma = 1$ and $A = \kappa \Delta$. Define $R_\lambda \colon \mathscr{S}_c(\mathscr{X}) \to \mathscr{S}_c(\mathscr{X})$ by the formula*

$$(R_\lambda f)_n := r_{\lambda, n}(f)$$
$$= \begin{cases} (\lambda + n(n-1) - A_n)^{-1} \sum_{m=n}^{\infty} \Pi^m_{n,\lambda} f_m & \text{if } n \geqslant 1 \\ \lambda^{-1} f_0 & \text{if } n = 0, \end{cases}$$

where $\Pi^n_{n,\lambda} \colon \mathscr{S}((\mathbb{R}^d)^m) \to \mathscr{S}((\mathbb{R}^d)^n)$ is the map defined by

$$\Pi^m_{n,\lambda} f_m = \left(\prod_{k=n+1}^{m} S^{k-1} (\lambda + k(k-1) - A_k)^{-1} \right) f_m,$$

if $m > n$ and $\Pi^m_{m,\lambda} = I$.

Then $(\lambda - \mathscr{L})\, \Lambda(R_\lambda f) = \Lambda(f)$ for all $f \in \mathscr{S}_c(\mathscr{X})$; hence, $\mathscr{R}_\lambda(f) = \Lambda(R_\lambda f)$.

The proposition gives us an explicit representation of the Fleming–Viot resolvent in terms of compositions of the operators S^{n-1} with the resolvent of the Laplacian in various dimensions.

Let us look more closely at a typical piece of this formula. Note that the operator $S^{n-1}_{ij} \colon \mathscr{S}((\mathbb{R}^d)^n) \to \mathscr{S}((\mathbb{R}^d)^{n-1})$ can be written as the restriction of a function on $(\mathbb{R}^d)^n$ to the hyperplane H^{n-1}_{ij} followed by a linear change of variable:

$$S^{n-1}_{ij} g_n = g_n \circ s^{n-1}_{ij} = g_n \restriction H^{n-1}_{ij} \circ s^{n-1}_{ij}.$$

Let us use the notation

$$\rho^{n-1}_{ij} \colon \mathscr{S}((\mathbb{R}^d)^n) \to \mathscr{S}(H^{n-1}_{ij})$$

for the operation of restriction and the notation

$$U^n_l = (l - \Delta_n)^{-1}$$

for the resolvent of the Laplacian in $(\mathbb{R}^d)^n$. Now we can write a typical piece of the resolvent formula in the proposition above as

$$S^{n-1}(\lambda+n(n-1)-\kappa\Delta_n)^{-1} f_n$$
$$=\kappa^{-1}\sum_{i=1}^{n}\sum_{i\neq j}^{n}(\rho_{ij}^{n-1}U^n_{\kappa^{-1}(\lambda+n(n-1))}f_n)\circ s_{ij}^{n-1};$$

that is, as a sum of a *resolvent of the Laplacian* followed by a *restriction to a hyperplane* followed by a *linear change of variable*. This observation is the key to our proof of the theorem. Indeed, each of these three types of operation is familiar and can be conveniently estimated in the scale of Sobolev spaces by well-known methods.

For $s\in\mathbb{R}$ and $n\geqslant 1$, let $H^s((\mathbb{R}^d)^n)$ denote the space of functions $f_n\in L^2(\mathbb{R}^d)^n)$ such that $(1+|\xi|^2)^{s/2}\hat{f}_n(\xi)\in L^2((\mathbb{R}^d)^n)$, where

$$\hat{f}_n(\xi)=(2\pi)^{-nd/2}\int_{(\mathbb{R}^d)^n} e^{-i(x,\xi)}f_n(x)\,dx$$

is the Fourier transform. Since the subscript of the function indicates the number of its $\mathbb{R}^d$-valued variables, the notation

$$\|f_n\|_s=\|f_n\|_{H^s((\mathbb{R}^d)^n)}=\left(\int_{(\mathbb{R}^d)^n}(1+|\xi|^2)^s\,|\hat{f}(\xi)|^2\,d\xi\right)^{1/2}$$

should cause no confusion.

LEMMA A. *Let* $U_l^n=(l-\Delta_n)^{-1}$. *Then for* $0\leqslant\beta\leqslant 2$,

$$\|U_l^n f_n\|_{s+\beta}\leqslant\left(\frac{1}{l}\right)^{1-\beta/2}\left(1+\frac{1}{l}\right)^{\beta/2}\|f_n\|_s.$$

Proof. If $f_n\in H^s((\mathbb{R}^d)^n)$, then

$$\begin{aligned}\|U_l^n f_n\|_s^2&=\int_{(\mathbb{R}^d)^n}|(1+|\xi|^2)^{s/2}(l+|\xi|^2)^{-1}\hat{f}_n(\xi)|^2\,d\xi\\&\leqslant\sup_\xi\left(\frac{1+|\xi|^2}{l+|\xi|^2}\right)^2\int_{(\mathbb{R}^d)^n}|(1+|\xi|^2)^{(s-2)/2}\hat{f}_n(\xi)|^2\,d\xi\\&\leqslant\left(1+\frac{1}{l}\right)^2\|f_n\|_{s-2}^2.\end{aligned}$$

Similarly, $\|U_l^n f_n\|_s^2\leqslant l^{-2}\|f_n\|_s^2$, and the result follows by interpolation. ∎

Next, let us consider the restriction operator ρ_{ij}^{n-1}. Let $(H_{ij}^{n-1})^\perp$ denote the orthogonal complement of H_{ij}^{n-1} in $(\mathbb{R}^d)^n$. It is a linear subspace of dimension d. Let us write

$$x=x'+x'',\qquad x'\in H_{ij}^{n-1},\qquad x''\in(H_{ij}^{n-1})^\perp$$

for this splitting, and also

$$\xi = \xi' + \xi''$$

for the dual splitting. Thus

$$(x, \xi) = (x', \xi') + (x'', \xi''),$$

since $(x', \xi'') = (x'', \xi') = 0$.

LEMMA B. *In the notation above,*

$$(\rho_{ij}^{n-1} f_n)^\wedge\,(\xi') = (2\pi)^{-d/2} \int_{[(H_{ij}^{n-1})^\perp]^*} \hat{f}_n(\xi' + \xi'')\, d\xi'',$$

and, for all $s \in \mathbb{R}$,

$$\|\rho_{ij}^{n-1} f_n\|_s^2 \leqslant \frac{\Gamma(s)}{(4\pi)^{d/2}\,\Gamma(s + d/2)}\, \|f_n\|_{s+d/2}.$$

Proof. Since, by Fourier inversion,

$$\rho_{ij}^{n-1} g_n(x') = (2\pi)^{-(n-1)\,d/2} \int_{(H_{ij}^{n-1})^*} e^{i(x', \xi')} (\rho_{ij}^{n-1} g_n)^\wedge\,(\xi')\, d\xi'$$

and also

$$\begin{aligned}
\rho_{ij}^{n-1} g_n(x') &= g_n(x') \\
&= (2\pi)^{-n(d/2)} \int_{(\mathbb{R}^d)^n} e^{i(x', \xi)} \hat{g}_n(\xi)\, d\xi \\
&= (2\pi)^{-(n-1)\,d/2} \int_{(H_{ij}^{n-1})^*} e^{i(x', \xi')} \\
&\qquad \times (2\pi)^{-d/2} \int_{[(H_{ij}^{n-1})^\perp]^*} \hat{g}_n(\xi' + \xi'')\, d\xi''\, d\xi',
\end{aligned}$$

It follows that

$$(\rho_{ij}^{n-1} g_n)^\wedge\,(\xi') = (2\pi)^{-d/2} \int_{[(H_{ij}^{n-1})^\perp]^*} \hat{g}_n(\xi' + \xi'')\, d\xi''.$$

The estimate for the H^s-norm of ρ_{ij}^{n-1} is well-known and follows by direct computation from the formula above as, for example, in [AS, Section 8, Theorem 1a]. ∎

Since $s_{ij}^{n-1}\colon (\mathbb{R}^d)^{n-1} \to H_{ij}^{n-1}$ is invertible we have

$$(s_{ij}^{n-1})^{-1}\colon H_{ij}^{n-1} \to (\mathbb{R}^d)^{n-1}$$

and if we make the usual identification of dual spaces $[(\mathbb{R}^d)^{n-1}]^* = (\mathbb{R}^d)^{n-1}$ and define

$$t_{ij}^{n-1} = [(s_{ij}^{n-1})^{-1}]^*,$$

then we have $t_{ij}^{n-1}\colon (\mathbb{R}^d)^{n-1} \to [H_{ij}^{n-1}]^*$. We use this notation in the following result.

Lemma C. *The linear transformation* $s_{ij}^{n-1}\colon (\mathbb{R}^d)^{n-1} \to H_{ij}^{n-1}$ *is invertible, with* $\|s_{ij}^{n-1}\| = \sqrt{2}$ *and* $|\det s_{ij}^{n-1}| = 2^{d/2}$. *Furthermore, if* $g_{n-1} \in \mathscr{S}(H_{ij}^{n-1})$

$$(g_{n-1} \circ s_{ij}^{n-1})^\wedge\,(\eta) = 2^{-d/2}\hat{g}_{n-1}(t_{ij}^{n-1}\eta).$$

Proof. Notice that s_{ij}^{n-1} is conformal: it takes orthogonal vectors to orthogonal vectors and leaves lengths unchanged in all directions except those in the $(i \wedge j)$th factor of $(\mathbb{R}^d)^{n-1}$. In those d directions lengths are stretched by a factor of $\sqrt{2}$. The first assertions of the lemma follow easily from these facts.

Now if $\eta \in [(\mathbb{R}^d)^{n-1}]^* = (\mathbb{R}^d)^{n-1}$, and we set $s = s_{ij}^{n-1}$, then

$$\begin{aligned}(g_{n-1} \circ s)^\wedge\,(\eta) &= (2\pi)^{-(n-1)\,d/2} \int_{(\mathbb{R}^d)^{n-1}} e^{-i(x,\,\eta)} g_{n-1}(sx)\,dx \\ &= (2\pi)^{-(n-1)\,d/2} \int_{H_{ij}^{n-1}} e^{-i(s^{-1}y,\,\eta)} g_{n-1}(y)\,|\det s^{-1}|\,dy \\ &= |\det s|^{-1}\,\hat{g}_{n-1}((s^{-1})^*\,\eta). \quad \blacksquare\end{aligned}$$

Proposition B. *For each* $n \geqslant 2$ *and each Schwarz class function* $f_n \in \mathscr{S}((\mathbb{R}^d)^n)$ *we have the formula*

$$\begin{aligned}&(S^{n-1}U_l^n g_n)^\wedge\,(\eta) \\ &\quad= (4\pi)^{-d/2} \sum_{i=1}^n \sum_{j \neq i}^n \int_{[(H_{ij}^{n-1})^\perp]^*} (l + |t_{ij}^{n-1}\eta + \zeta|^2)^{-1}\,\hat{g}_n(t_{ij}^{n-1}\eta + \zeta)\,d\zeta.\end{aligned}$$

Consequently, for each $s > 0$,

$$\begin{aligned}&\|S^{n-1}U_l^n f_n\|^2_{H^s((\mathbb{R}^d)^{n-1})} \\ &\quad\leqslant \frac{2^s}{(16\pi)^{d/2}} \left(\frac{n(n-1)}{l}\right)^2 \frac{\Gamma(s)}{\Gamma(s+d/2)}\,\|f_n\|^2_{H^{s+d/2}((\mathbb{R}^d)^n)}.\end{aligned}$$

Proof. By Lemmas B and C,

$$\begin{aligned}
&(S^{n-1}U^n_l f_n)^\wedge(\eta)\\
&\quad=\sum_{i=1}^{n}\sum_{j\neq i}^{n}[(\rho^{n-1}_{ij}U^n_l f_n)\circ s^{n-1}_{ij}]^\wedge(\eta)\\
&\quad=2^{-d/2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}(\rho^{n-1}_{ij}U^n_l f_n)^\wedge(t^{n-1}_{ij}\eta)\\
&\quad=2^{-d/2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}(2\pi)^{-d/2}\int_{[(H^{n-1}_{ij})^\perp]^*}(U^n_l f_n)^\wedge(t^{n-1}_{ij}\eta+\zeta)\,d\zeta\\
&\quad=(4\pi)^{-d/2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\int_{[(H^{n-1}_{ij})^\perp]^*}(l+|t^{n-1}_{ij}\eta+\zeta|^2)^{-1}\hat{f}_n(t^{n-1}_{ij}\eta+\zeta)\,d\xi.
\end{aligned}$$

Now set $M^n_l = lU^n_l$. By Lemma A, M^n_l is a bounded operator of norm 1 on $H^s((\mathbb{R}^d)^n)$ for all s. Thus,

$$\begin{aligned}
&\|S^{n-1}U^n_l f_n\|^2_{H^s((\mathbb{R}^d)^{n-1})}\\
&\quad=\int_{(\mathbb{R}^d)^{n-1}}(1+|\eta|^2)^s\,|(S^{n-1}U^n_l f_n)^\wedge(\eta)|^2\,d\eta\\
&\quad=2^{-d}\left(\frac{n(n-1)}{l}\right)^2\int_{(\mathbb{R}^d)^{n-1}}(1+|\eta|^2)^s\\
&\qquad\times\left|\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}^{n}(\rho^{n-1}_{ij}M^n_l f_n)^\wedge(t^{n-1}_{ij}\eta)\right|^2 d\eta\\
&\quad\leqslant 2^{-d}\left(\frac{n(n-1)}{l}\right)^2\frac{1}{n(n-1)}\\
&\qquad\times\sum_{i=1}^{n}\sum_{j\neq i}^{n}\int_{(\mathbb{R}^d)^{n-1}}(1+|\eta|^2)^s\,|(\rho^{n-1}_{ij}M^n_l f_n)^\wedge(t^{n-1}_{ij}\eta)|^2\,d\eta.
\end{aligned}$$

Now since $(t^{n-1}_{ij})^{-1}=(s^{n-1}_{ij})^*$, it follows that

$$\begin{aligned}
&\int_{(\mathbb{R}^d)^{n-1}}(1+|\eta|^2)^s\,|(\rho^{n-1}_{ij}M^n_l f_n)^\wedge(t^{n-1}_{ij}\eta)|^2\,d\eta\\
&\quad=\int_{(H^{n-1}_{ij})^*}(1+|(s^{n-1}_{ij})^*\,\xi'|^2)^s\,|(\rho^{n-1}_{ij}M^n_l f_n)^\wedge(\xi')|^2\,d\xi'\\
&\quad=\int_{(H^{n-1}_{ij})^*}\left(\frac{1+|(s^{n-1}_{ij})^*\,\xi'|^2}{1+|\xi'|^2}\right)^s(1+|\xi'|^2)^s\,|(\rho^{n-1}_{ij}M^n_l f_n)^\wedge(\xi')|^2\,d\xi'
\end{aligned}$$

$$\leqslant \|s_{ij}^{n-1}\|^{2s} \|\rho_{ij}^{n-1} M_l^n f_n\|_{H^s(H_{ij}^{n-1})}^2$$
$$\leqslant \frac{2^s \Gamma(s)}{(4\pi)^{d/2}\, \Gamma(s+d/2)} \|f_n\|_{H^{s+d/2}((\mathbb{R}^d)^n)}^2.$$

Combining this calculation with the one above finishes the proposition. ∎

Proof of the Theorem. First let us verify that $\|\cdot\|_{\alpha,\beta}$ is indeed a norm on polynomials finer than the supremum norm. In fact, reverting to the notation of the introduction, if

$$\|f_n\|_{C(X)} \leqslant C_n \|f_n\|_{(n)} \qquad \text{and} \qquad \sup_n \frac{C_n}{\rho_n} = M < \infty,$$

then $\|\cdot\|_\rho$ defines a norm on polynomials finer than the sup norm. That it is a semi-norm is easy to see, so suppose $\|F\|_\rho = 0$. Then there are $f^{(N)} \in \mathscr{S}_c(\mathscr{X})$ such that $F = \Lambda(f^{(N)})$ and $\sum_{n=0}^\infty \rho_n \|f_n^{(N)}\|_{(n)} \to 0$ as $N \to \infty$. But then,

$$0 \leqslant |F(\mu)| \leqslant \sum_{n=0}^\infty |\langle f_n^{(N)}, \mu^n\rangle| \leqslant \sum_{n=0}^\infty C_n \|f_n^{(N)}\|_{(n)}$$
$$\leqslant M \sum_{n=0}^\infty \rho_n \|f_n^{(N)}\|_{(n)};$$

hence F is identical, zero.

In our case $\sup_n C_n < \infty$ and the weights ρ_n increase without bound so the condition above is trivially satisfied. To be precise, let $s \geqslant v_n = 1 + nd/2$ and let $f_n \in \mathscr{S}((\mathbb{R}^d)^n)$. Then, imitating the proof of Sobolev's lemma,

$$\sup_x |f_n(x)| = \sup_x \left| (2\pi)^{-nd/2} \int_{(\mathbb{R}^d)^n} e^{i(x,\xi)} \hat{f}_n(\xi)\, d\xi \right|$$
$$\leqslant (2\pi)^{-nd/2} \int_{(\mathbb{R}^d)^n} (1+|\xi|^2)^{s/2} |\hat{f}_n(\xi)| (1+|\xi|^2)^{-s/2}\, d\xi$$
$$\leqslant (2\pi)^{-nd/2} \|f_n\|_s \times \left(\int_{(\mathbb{R}^d)^n} (1+|\xi|^2)^{-s}\, d\xi \right)^{1/2}$$
$$= (2\pi)^{-nd/2} \|f_n\|_s \times \left(\frac{2\pi^{nd/2}}{\Gamma(nd/2)} \int_0^\infty (1+r^2)^{-s} r^{nd-1}\, dr \right)^{1/2}.$$

It is easy to check that the constant above is bounded in n uniformly in $s \geqslant v_n$, and that does the trick.

Next, let us note that by scaling, we need only prove the resolvent estimates for $\gamma = 1$ and general $\kappa > 0$. For if $\mathscr{L}$ is the Fleming–Viot operator with drift and diffusion coefficients $\kappa\Delta$ and γ, respectively, and $\mathscr{L}'$ is the Fleming–Viot operator with drift and diffusion coefficients $(\kappa/\gamma)\,\Delta$ and 1, respectively, then $\mathscr{L}=\gamma\mathscr{L}'$. Hence,

$$\mathscr{R}_\lambda=(\lambda-\mathscr{L})^{-1}=\gamma^{-1}\left(\frac{\lambda}{\gamma}-\mathscr{L}'\right)^{-1}.$$

So now we assume $\gamma=1$ and $\kappa>0$. By Proposition B with $s=\nu_{n-1}$,

$$\begin{aligned}
&\|S^{n-1}(\lambda+n(n-1)-\kappa\Delta_n)^{-1}\,f_n\|_{\nu_{n-1}}\\
&\quad=\kappa^{-1}\,\|S^{n-1}U^n_{\kappa^{-1}(\lambda+n(n-1))}\,f_n\|_{\nu_{n-1}}\\
&\quad\leqslant\kappa^{-1}\,\frac{n(n-1)}{\kappa^{-1}(\lambda+n(n-1))}\sqrt{\frac{2^{\nu_{n-1}}\Gamma(\nu_{n-1})}{(16\pi)^{d/2}\,\Gamma(\nu_n)}}\,\|f_n\|_{\nu_n}\\
&\quad\leqslant\frac{\omega_n}{\omega_{n-1}}\,\|f_n\|_{\nu_n}.
\end{aligned}$$

Since

$$\Pi^m_{n,\lambda}\,f_m=S^n(\lambda+(n+1)\,n-\kappa\Delta_{n+1})^{-1}\,\Pi^m_{n+1,\lambda}\,f_m,$$

iterating this estimate gives

$$\|\Pi^m_{n,\lambda}\,f_m\|_{\nu_n}\leqslant\frac{\omega_m}{\omega_n}\,\|f_m\|_{\nu_m}.$$

We come now to the proof of item (ii) of the theorem. The proof of item (i), being so similar, is left to the reader. So suppose $0\leqslant\beta\leqslant 2$ and $f\in\mathscr{S}_c(\mathscr{X})$. By Lemma A we have

$$\begin{aligned}
&\|r_{\lambda,n}(f)\|_{\nu_n+\beta}\\
&\quad=\kappa^{-1}\left\|U^n_{\kappa^{-1}(\lambda+n(n-1))}\sum_{m=n}^{\infty}\Pi^m_{n,\lambda}\,f_m\right\|_{\nu_n+\beta}\\
&\quad\leqslant\kappa^{-1}\left(\frac{\kappa}{\lambda+n(n-1)}\right)^{1-\beta/2}\left(1+\frac{\kappa}{\lambda+n(n-1)}\right)^{\beta/2}\left\|\sum_{m=n}^{\infty}\Pi^m_{n,\lambda}\,f_m\right\|_{\nu_n}\\
&\quad\leqslant\left(\frac{1}{\lambda+n(n-1)}\right)^{1-\beta/2}\left(\frac{1}{\kappa}+\frac{1}{\lambda}\right)^{\beta/2}\sum_{m=n}^{\infty}\frac{\omega_m}{\omega_n}\,\|f_m\|_{\nu_m}.
\end{aligned}$$

Thus, if $\alpha+\beta>1$,

$$\sum_{n=0}^{\infty} (1+n)^{\alpha} \omega_n \|r_{\lambda, n}(f)\|_{v_n+\beta}$$

$$\leqslant (\kappa^{-1}+\lambda^{-1})^{\beta/2} \sup_n \left(\frac{(1+n)^2}{\lambda+n(n-1)}\right)^{1-\beta/2}$$

$$\times \sum_{n=0}^{\infty} (1+n)^{\alpha+\beta-2} \sum_{m=n}^{\infty} \omega_m \|f_m\|_{v_m}$$

$$\leqslant (\kappa^{-1}+\lambda^{-1})^{\beta/2} (1+\lambda^{-1})^{1-\beta/2} \sum_{m=0}^{\infty} \omega_m \|f_m\|_{v_m} \times C \frac{(1+m)^{\alpha+\beta-1}}{\alpha+\beta-1}.$$

Now $\Lambda(f)$ is a representation of some $F \in P^{\mathscr{S}}(\mathscr{M}_1(\mathbb{R}^d))$. Taking first the infimum on the left over all such representations, then the corresponding infimum on the right, we obtain

$$\|\mathscr{R}_\lambda F\|_{\alpha, \beta} \leqslant C \frac{(1+\lambda^{-1})^{1-\beta/2} (\kappa^{-1}+\lambda^{-1})^{\beta/2}}{\alpha+\beta-1} \|F\|_{\alpha+\beta-1, 0},$$

which was to be proved.

REFERENCES

[AS] N. Aronszajn and K. T. Smith, Theory of Bessel potentials: Part I, *Ann. Inst. Fourier, Grenoble* (1961), 385–475.

[D] D. A. Dawson, Measure-valued Markov processes, *in* "Ecole d'Eté Probabilités de Saint-Flour XXI, 1991," Lecture Notes in Mathematics, Vol. 1541, Springer-Verlag, Berlin/New York, 1993.

[DM] D. A. Dawson and P. March, Resolvent estimates for Fleming–Viot operators and uniqueness of solutions to related martingale problems, *J. Funct. Anal.* **132** (1995), 417–472.

[DK] P. Donnelly and T. G. Kurtz, A countable representation of the Fleming–Viot measure-valued diffusion, *Ann. Probab.* **24** (1996), 698–742.

[E] S. N. Ethier, A class of degenerate diffusion processes occurring in population genetics, *Comm. Pure Appl. Math.* **29** (1976), 483–493.

[EK1] S. N. Ethier and T. G. Kurtz, "Markov Processes: Characterization and Convergence," Wiley, New York, 1985.

[EK2] S. N. Ethier and T. G. Kurtz, Fleming–Viot processes in population genetics, *SIAM J. Cont. Opt.* **31** (1993), 345–386.

[F] W. Feller, Diffusion Processes in Genetics, *in* "Second Berkeley Symposium on Mathematical Statistics and Probability," pp. 227–246, University of California Press, Berkeley, 1951.

[FV] W. H. Fleming and M. Viot, Some measure-valued processes in population genetics theory, *Indiana Univ. Math. J.* **28** (1979), 817–843.

[Sa] K. Sato, Diffusion processes and a class of Markov chains related to population genetics, *Osaka J. Math.* **13** (1976), 631–659.

[Sh] N. Shimakura, Equations différentielles provenant de la génétique des populations, *Tohoku Math. J.* **29** (1977), 287–318.